551.[58]:63:001.5

SECTION II.—GENERAL METEOROLOGY.

THE AVERAGE INTERVAL CURVE AND ITS APPLICATION TO METEOROLOGICAL PHENOMENA.[1]

By W. J. SPILLMAN, H. R. TOLLEY, and W. G. REED.

[Dated: U. S. Office of Farm Management, Washington, Apr. 13, 1916.]

In cooperation with the Bureau of Plant Industry the Office of Farm Management recently undertook a study of methods and cost of heating greenhouses. In these investigations detailed studies were made of the heating systems in a large number of commercial establishments. In prosecuting this investigation in central and southern latitudes we met everywhere the question on the part of the growers: "What is the lowest temperature we are liable to have once in 30 years?" These men were of opinion that it would be cheaper to stand a loss once on 30 years than to go to the expense of providing against it.

Temperature is, of course, not the only factor in the problem. Wind velocity is quite as important. It is the combination of low temperature and high wind that sends the cold chills down the back of the man who has extensive investments in greenhouses.

It may be of interest to note that we did not meet the above query when these investigations were extended into the northern districts. Greenhouse men there informed us that when the temperature falls below about 20° F. the moisture in the house freezes on the glass and seals up the chinks, so that it actually costs less to heat a hothouse full of growing plants with outside temperatures below 20° than with temperatures slightly above this point.

It is not difficult when the mean and the standard deviation of the extreme winter minima are known for a given locality, to calculate the average frequency with which temperatures lower than any assigned limit will occur or the temperature limit that will be exceeded once on the average in a given number of years. However, the labor involved is considerable, especially when the calculations must be made for several stations, and more especially when several temperatures or intervals are concerned. In making these calculations for a large number of stations it occurred to the senior author of this paper to shorten the labor by constructing a curve with average intervals between successive occurrences of a minimum lower than a given temperature, T, as ordinates and with the departure of T from the mean as abscissæ, T being expressed in terms of the standard deviation as the unit. Accordingly he constructed such a curve (see fig. 1). The equation of this curve was derived from the following considerations: In the theory of probability, unity represents certainty. An even chance is represented by the fraction $\frac{1}{2}$. Assuming that the frequency curve for the extreme winter minima is normal, which it is nearly, it is an even chance whether the lowest temperature during any one winter will be above or below its average value for any particular locality. The probability that it will be below the mean is therefore $\frac{1}{2}$, as indicated in figure 2. Let P represent the probability that it will lie between the mean, M, and some temperature, T, lower than the mean. Then the probability that it will lie below T is $\frac{1}{2}-P$. Suppose that in a given case the value of this latter probability is 1/10. This means that the chance is one in ten that the minimum for the winter will be lower than T, or that the minimum will be

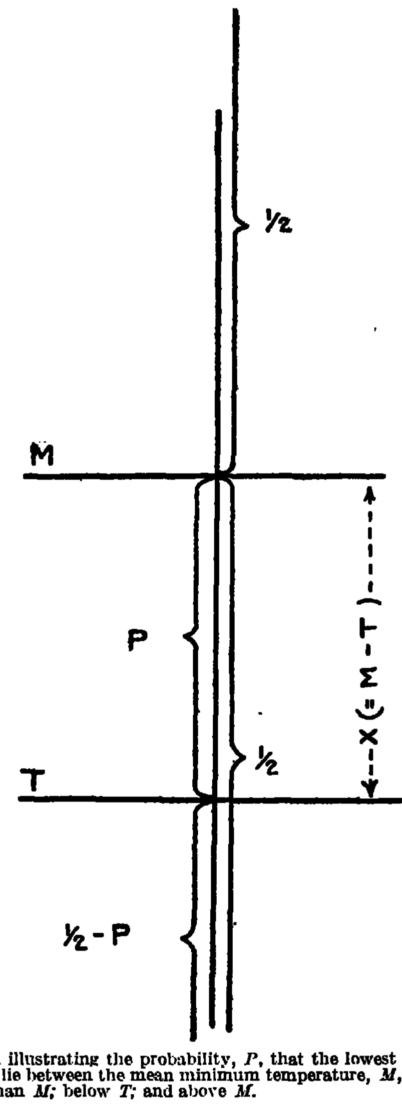


FIG. 2.—Diagram illustrating the probability, P, that the lowest temperature of any one winter will lie between the mean minimum temperature, M, and some temperature, T, lower than M; below T; and above M.

lower than T on the average once in ten years. Thus, the probability is 1/10, and the average interval is 10. The average interval between the occurrence of minima below T is thus the reciprocal of the probability of such occurrence, a general principle, applicable to all cases. If, then, the average interval be represented by A, we have $A=1/(\frac{1}{2}-P)$. The value of P can be expressed in terms of the standard deviation and the departure of T from the

[1] A paper read before the Philosophical Society of Washington, Apr. 1, 1916.

mean, which departure is usually represented by X. Substituting this value for P in the last equation we have

$$A=\frac{1}{\frac{1}{2}-\frac{1}{\sqrt{2\pi}}\left(\frac{X}{D}-\frac{1}{3\times 2}\frac{X^3}{D^3}+\frac{1}{\underline{|2}\times 5\times 4}\frac{X^5}{D^5}-\frac{1}{\underline{|3}\times 7\times 8}\frac{X^7}{D^7}+\&c.\right)}.$$

The above equation is the equation of the average interval curve, the values of A constituting the ordinates and those of X/D the abscissæ.

Table 1 has been constructed to give a comparison of the use of this curve as compared with the use of the ordinary probability integral tables in problems of the kind here under consideration. Suppose the problem is to find the average interval between successive occurrences of an absolute winter minimum lower than T.

TABLE 1.—*Comparison of use of tables and the curve.*

Given M, D, and T to find A.

TABLES	CURVE
1. $M-T=X$ 2. X/D 3. Value of P corresponding to X/D 4. $A=1/(\frac{1}{2}-P)$	1. $M-T=X$ 2. X/D 3. Ordinate A corresponding to X/D

Given M, D, and A to find T.

1. $P=(A-2)/2A$ 2. Value of X/D corresponding to P 3. $(X/D)D=X$ 4. $T=M-X$	1. Abscissa X/D corresponding to A 2. $(X/D)D=X$ 3. $T=M-X$

We first calculate the mean and the standard departure of the extreme winter minima, as shown in figure 1. To find A by means of the tables we then subtract T from M to obtain X; divide X by the standard deviation; then from the tables we find the value of P corresponding to this value of X/D. Finally we subtract the value of P from $\frac{1}{2}$ and divide the result into unity.

In using the curve the first two operations are the same. Having found X/D, we find the ordinate corresponding to abscissa X/D. This ordinate is the average interval sought. The use of the curve thus saves one complex operation, and substitutes the reading of a curve for hunting in a table for a number after the manner of finding the number corresponding to a given logarithm.

If A be given and it is desired to find T, which was the problem given us by the greenhouse men, the operations are as shown in the lower part of Table 1. Here again we save one operation still more complex, and substitute the reading on the curve for the search in a table of figures.

We shall first illustrate the use of the curve in some concrete examples, and incidentally show something of the applicability of the method to meteorological prediction.

The first column of Table 2 gives the absolute winter minimum at Philadelphia for 24 successive winters. This number of values of the variable is too small for any great degree of accuracy, but will serve for purposes of illustration. We first calculate the mean and the standard deviation in the usual way (see fig. 1). It will be noticed that 13 of the minima of Table 2, or one more than half, are below the mean, which is 5.04°F. This gives some idea of the closeness of fit of the normal frequency curve in this case, since theoretically half of them should be above and half below the mean.

TABLE 2.—*Absolute winter minimum temperatures at Philadelphia, and calculation of the mean and the standard deviation.*

Minima.	X'.	$(X')^2$.	Calculation.
°F.			
19	14	196	
4	1	1	
10	5	25	
−5	10	100	
4	1	1	
8	3	9	$M=121/24=5.04$.
12	7	49	
3	2	4	$M'=5$.
−5	10	100	
−1	6	36	$n=24$.
6	1	1	
9	4	16	
10	5	25	
5	0	0	
−2	7	49	$D=\sqrt{\frac{S(X'^2)}{n}-\frac{r^2}{n^2}}=5.877$.*
7	2	4	
2	3	9	(For meaning of symbols see explanations in fig. 1, facing p. 198.)
2	3	9	
9	4	16	
13	8	64	
10	5	25	
0	5	25	
4	1	1	
−3	8	64	
Totals... 121	....	829	

* This modification of the common formula for computing the standard deviation, is due to Prof. C. F. Marvin, Chief of the Weather Bureau, who has kindly consented to its use here in advance of publication. The new formula is much more convenient for computation than the one commonly used.—*W. J. S.*

Let us now find the temperature below which the winter minimum should be expected to fall once on the average in 30 years. The abscissa corresponding to ordinate 30 on the curve (fig. 1) is 1.834. This, then, is the value of X/D. Multiplying by 5.877, the previously found value of D, we find $X=10.778$, which gives for T the value of -5.738 degrees (F.). Theoretically, the lowest winter temperature at Philadelphia should thus fall below $-5.738°$ once on the average in 30 years. Reference to Table 2 shows that during the 24-year period of observation there recorded the temperature did not quite reach this low level, but did reach $-5°$ on two occasions.

To further test the reliability of this method of prediction let us determine how often the temperature should fall below zero. In this case $X=5$, and $X/D=0.85$. On the curve this abscissa corresponds to $A=5$; hence the temperature should fall below 0°F. once on the average in five years. During the 24 years for which data are available it did fall below five times, which agrees with the calculated value of A.

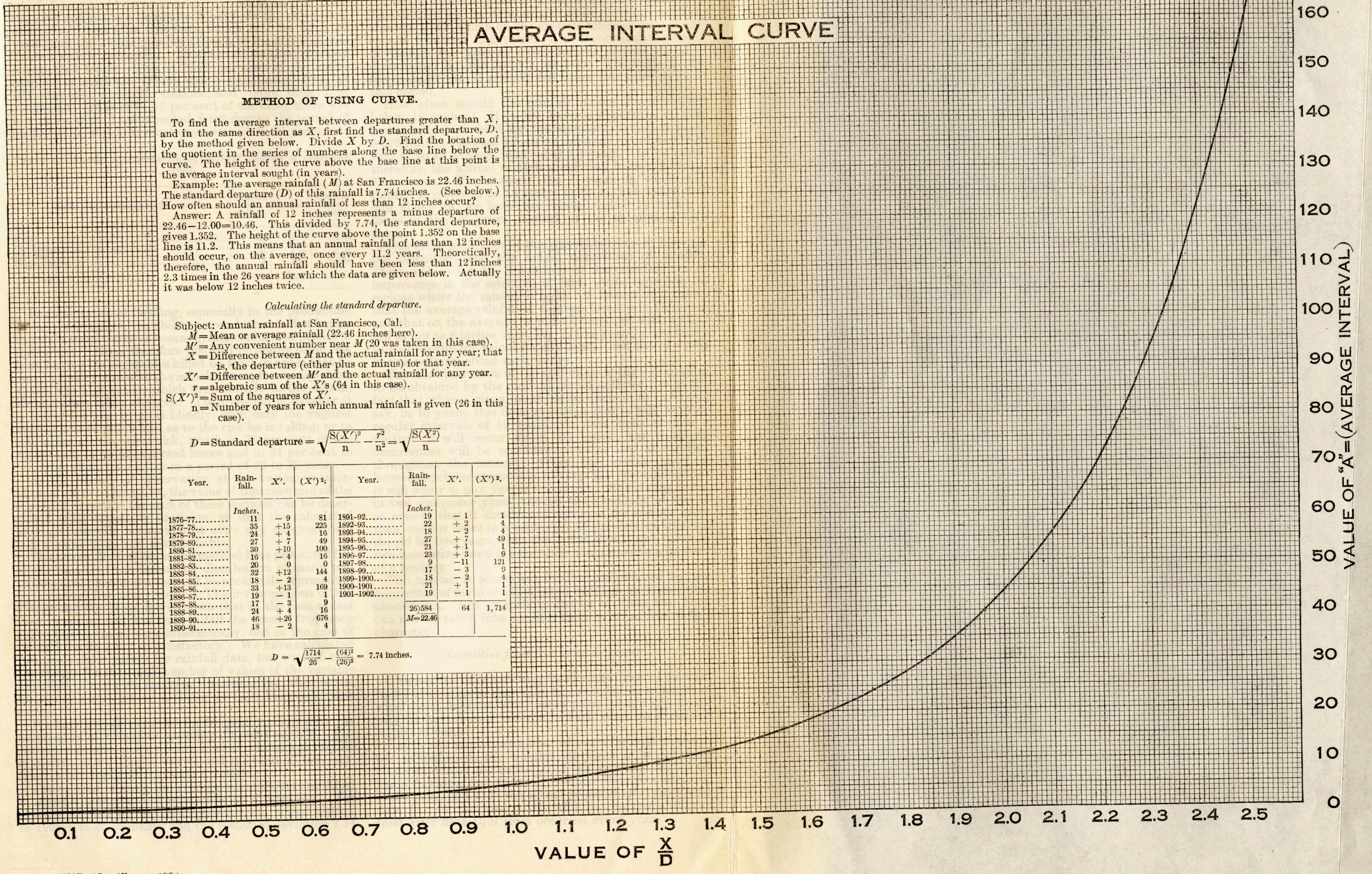
There are three important limitations in the application of the method here outlined to meteorological phenomena. The first lies in the fact that the periods for which data are available give so few values of the variable that the resulting values of the mean and the standard deviation are not as reliable as it is desirable they should be. Nevertheless, when the frequency curve is approximately normal, the results calculated from 20 years' observations are fairly satisfactory, as will be seen below.

The two junior authors of this paper have calculated the standard deviation of the date of last killing frost in spring for 569 stations, the data for which were kindly furnished by the Weather Bureau. The length of the record at these stations covers 10 to 59 years, as shown in Table 3.

TABLE 3.—*Records used as a test of the reliability of the method.*

Length of record (years).	Number of stations.
10–16	6
17–19	43
20–23	338
24–29	115
30–39	43
40–46	23
59	1
Total	569

AVERAGE INTERVAL CURVE



METHOD OF USING CURVE.

To find the average interval between departures greater than X, and in the same direction as X, first find the standard departure, D, by the method given below. Divide X by D. Find the location of the quotient in the series of numbers along the base line below the curve. The height of the curve above the base line at this point is the average interval sought (in years).

Example: The average rainfall (M) at San Francisco is 22.46 inches. The standard departure (D) of this rainfall is 7.74 inches. (See below.) How often should an annual rainfall of less than 12 inches occur?

Answer: A rainfall of 12 inches represents a minus departure of 22.46−12.00=10.46. This divided by 7.74, the standard departure, gives 1.352. The height of the curve above the point 1.352 on the base line is 11.2. This means that an annual rainfall of less than 12 inches should occur, on the average, once every 11.2 years. Theoretically, therefore, the annual rainfall should have been less than 12 inches 2.3 times in the 26 years for which the data are given below. Actually it was below 12 inches twice.

Calculating the standard departure.

Subject: Annual rainfall at San Francisco, Cal.

M = Mean or average rainfall (22.46 inches here).

M' = Any convenient number near M (20 was taken in this case).

X = Difference between M and the actual rainfall for any year; that is, the departure (either plus or minus) for that year.

X' = Difference between M' and the actual rainfall for any year.

r = algebraic sum of the X's (64 in this case).

$S(X')^2$ = Sum of the squares of X'.

n = Number of years for which annual rainfall is given (26 in this case).

$$D = \text{Standard departure} = \sqrt{\frac{S(X')^2}{n} - \frac{r^2}{n^2}} = \sqrt{\frac{S(X^2)}{n}}$$

Year.	Rainfall.	X'.	$(X')^2$.	Year.	Rainfall.	X'.	$(X')^2$.
	Inches.				*Inches.*		
1876-77	11	−9	81	1891-92	19	−1	1
1877-78	35	+15	225	1892-93	22	+2	4
1878-79	24	+4	16	1893-94	18	−2	4
1879-80	27	+7	49	1894-95	27	+7	49
1880-81	30	+10	100	1895-96	21	+1	1
1881-82	16	−4	16	1896-97	23	+3	9
1882-83	20	0	0	1897-98	9	−11	121
1883-84	32	+12	144	1898-99	17	−3	9
1884-85	18	−2	4	1899-1900	18	−2	4
1885-86	33	+13	169	1900-1901	21	+1	1
1886-87	19	−1	1	1901-1902	19	−1	1
1887-88	17	−3	9		26)584	64	1,714
1888-89	24	+4	16		M=22.46		
1889-90	46	+26	676				
1890-91	18	−2	4				

$$D = \sqrt{\frac{1714}{26} - \frac{(64)^2}{(26)^2}} = 7.74 \text{ inches.}$$

Mr. Tolley and Mr. Reed have also calculated, by means of the average interval curve, the theoretical date after which frost should occur once in 10 years at each of these stations, and made a comparison between the theoretical number of frosts after this date, which is 1 for each 10 years, and the actual number at each of the 569 stations. Table 4 shows the results. At 414 stations there were no unexpected frosts. This is 72.75 per cent of the entire number of stations. There was 1 unexpected frost at 123 stations, 2 at 29 stations, and more than 2 at 3 stations. Thus in 94.35 per cent of the cases there was not more than 1 unexpected frost during the period of observation, which in most cases was 20 to 29 years.

TABLE 4.—*Number of stations having specified numbers of unexpected frosts.*

Stations having unexpected frosts.	Number of stations.	Per cent.	Cumulative per cent.
None	414	72.75	72.75
One frost	123	21.60	94.35
Two frosts	29	5.10	99.45
Three or more	3	0.55	100.00
Sums	569	100.00	....

In some types of farming, especially in the production of early vegetables, the ability to reach the market before the main supply arrives means greatly enhanced profits to the farmer. He is therefore justified in taking some risk from having his vegetables killed by frost in order that he may in those years when he is lucky enough to escape frost obtain the high prices which prevail early in the season. He therefore takes chances. The question is whether it is better for him to go it blindly, with only the vaguest impressions to guide him as to the risk he is taking, or to rely upon a theoretical risk, which in 73 per cent of the cases leads to no unexpected losses and in 94 per cent to not more than 1 such loss in a period of 20 to 30 years.

As the periods of observation at the various meteorological stations lengthen, the value of the mean and of the standard deviation for each of these stations can be determined with increasing precision, so that the calculated average intervals between frost after any assigned date will become more and more reliable.

The second limitation of the method here presented lies in the fact that the frequency distribution of many meteorological phenomena can not be accurately represented by a normal frequency curve. The investigations of Mr. Tolley and Mr. Reed have shown that in the case of last frost in spring and first frost in fall the normal frequency curve fits the facts very satisfactorily, but in the case of rainfall the fit is not so satisfactory. We have calculated Pearson's skew curves for rainfall data, but we find that the results for the small number of values of the variable available in any case are very unsatisfactory, the normal frequency curve giving a better fit than Pearson's skew curves.

We have also constructed the frequency polygon for the occurrence of droughts covering different periods. The corresponding frequency curve is what is known to mathematicians as a J-shaped curve, and the method here outlined is not applicable to such a variable.

The method of using the average interval curve outlined in figure 1 makes use of rainfall as an illustration. The average rainfall at San Francisco for the period 1876–7 to 1901–2 is 22.46 inches and the standard deviation 7.74 inches. Under these conditions, how often should the rainfall be less than 12 inches?

Twelve inches represents a departure of 10.46 inches. This divided by the standard deviation, 7.74 inches, gives abscissa 1.352, which on the average interval curve corresponds to about 11.2 years. The theoretical frequency of rainfalls less than 12 inches for the 26-year period of observation is thus 26/11.2, or 2.32. The table in the figure shows that it actually fell below this value twice in this period.

How often should the rainfall at San Francisco fall below 18 inches? Here we have a departure of 4.46 inches, which divided by the standard deviation gives 0.576 for the abscissa, which corresponds to ordinate 3.6; that is, to an average interval of 3.6 years. In 26 years, therefore, the annual rainfall should fall below 18 inches 7.22 times. If we count the 4 years in which the rainfall is recorded as 18 inches as being half below and half above 18, we have the rainfall below 18 inches 6 times in the 26 years, instead of the theoretical 7.22 times.

This gives some idea of the general agreement between the calculated and the actual departures of rainfall. These values for the frequency of low rainfall are of importance in the selection of farm lands, especially in regions where the rainfall is light. A farmer who knows that the average rainfall in a given locality is 15 inches and that on the average he must be content with a rainfall as low as 8 inches once, say, in five years, is less likely to pay exorbitant prices for dry lands than the one who knows nothing about such things.

The third limitation of the usefulness of the information to be obtained by the method herein outlined lies in the fact that when a given event occurs once on the average, say, in 10 years, this does not mean that it will occur at regular intervals of 10 years. It does mean that in a century it will occur about 10 times; but these 10 occurrences will be scattered more or less at random throughout the century. When the average interval between such occurrences is 10 years the probability that it will occur in any year is one-tenth. But it may occur in two successive years, though the likelihood of this is so remote that it should not occur *on the average* more than once or twice in a century. If the average interval should be five years, then the event should not occur in two successive years oftener than about once in a quarter of a century.

The matter is somewhat further complicated by the occurrence of more or less distinct cycles in meteorological phenomena. But so little is known of these that they can not be taken into consideration in an article of this character.

Localities frost free in some years.

The method of calculating the average interval between frosts after a given date may be extended to localities in which frost does not occur every year, though it is necessary to have longer records in such cases to make the results of value.

Let L represent the number of years for which records are available, R the number of years in which frost occurred during the period L, A the average interval (in years) between last frosts after date C if frost occurred every year, and $1/Y$ the actual probability of frost after date C. In this case $1/A$ would be the probability of frost after date C, if frost occurred every year.

Case I. When the probability of frost after date C is less than half the total probability of frost; that is, when $1/Y$ is less than half of R/Y. The value of A in this case is found from the proportion

$$1/Y : 1/A :: R : L,$$

from which $A = RY/L$. Since R and L are known, either A or Y can be found when the other is known, and one of them is always known in problems of the kind here under consideration.

Case II. When the probability of frost after date C is just half the total probability of frost; that is, when $1/Y = R/2L$. Substituting this value of $1/Y$ in the proportion of Case I, and reducing, we have $A = 2$. In this case the date C coincides with the average date of last frost.

Case III. When the probability of frost after date C is greater than half of R/L and less than R/L. In this case the value of A is less than 2, and is hence not on the curve; but if we deal with the probability of last frost occurring before date C, instead of after C, we shall then find a value of A that does lie on the curve. The date C will then come before the mean date of last frost.

Since the total probability of frost is R/L, the probability of last frost occurring before date C is $R/L - 1/Y$, or $(RY - L)/LY$. Hence we have

$$(RY - L)/LY : 1/A :: R : L,$$

from which $A = RY/(RY - L)$.

Case IV. When the probability of frost after date C is equal to the total probability of frost, or $1/Y = R/L$. Here A is infinite, and there is no solution.

Case V. When the probability of frost after date C is greater than the total probability of frost, or when $1/Y$ is greater than R/L. In this case A is negative and there is no solution.

5 51.578.1 (1/4)

A CORRELATION BETWEEN THE RAINFALL OF NORTH AND SOUTH AMERICA.

By H. HELM CLAYTON.

[Dated: Oficina Meteorológica Argentina, Buenos Aires, Mar. 20, 1916.]

Pursuing a line of research outlined in the Popular Science Monthly (New York) of December, 1901, the writer obtained the average rainfall in the United States between the longitudes of 80° W. and 110° W., which includes all the States except the north Atlantic, the Plateau, and the Pacific coast. This average was compiled from the data published in the bulletin of the Weather Bureau entitled "The Annual Precipitation of the United States for the Years 1872 to 1907," in which the rainfall is given for selected stations nearly equally distributed. Mr. P. C. Day has kindly extended this data to the end of 1914.

The object of the research was to compare the total rainfall with the total crop production, and between the two there is an interesting correlation. There has also been discovered a correlation between this rainfall of central North America and that of central South America as indicated by the outflow in the River Paraná. There appears to be an inverse correlation between these two and the rainfall of Australia. The data are given in Table 1.

TABLE 1.—*Comparison of annual rainfall over the central United States with that over Australia and the mean annual heights of the Paraná at Rosario, Argentina.*

Year.	Mean general rainfall in the United States between meridians 80° and 110° W.		Paraná at Rosario.		Australia.
	Annual fall.	Departure from mean.	Mean annual height.*	Departure from mean.	Percentage of area with rainfall above the average.
	Inches.	*Inches.*	*Meters.*	*Meters.*	*Per cent.*
1900	31.97	+2.11	4.527	+0.792	....
1901	26.75	−3.11	2.998	−0.737	....
1902	30.44	+0.58	3.536	−2.001	....
1903	30.14	+0.28	3.268	−0.467	....
1904	28.08	−1.78	3.807	+0.072	....
1905	34.18	+4.32	5.611	+1.876	....
1906	32.71	+2.85	3.621	−0.114	....
1907	28.77	−1.09	3.634	−0.101	....
1908	30.93	+1.07	4.249	+0.514	33
1909	30.09	+0.23	2.924	−0.811	40
1910	23.63	−6.23	2.837	−0.898	75
1911	28.37	−1.49	3.128	−0.607	25
1912	31.67	+1.81	4.382	+0.647	12
1913	30.80	+0.94	3.664	−0.071	27
1914	29.34	−0.53	3.836	+0.101	11
Mean	29.86.	....	3.735	....	....

* From data kindly furnished me by the Chief of the Hydrometric Section of the Oficina Meteorológica Argentina.

In this table column 1 gives the year; column 2 gives the mean annual rainfall per station in the United States; column 3 gives the departures from the average values; column 4 gives the mean annual river stages at Rosario; column 5 the departures of these from the average; column 6 gives the percentage of the area of Australia over which the rainfall was above the normal. These percentages were taken from a meteorological chart published by Mr. H. A. Hunt, Commonwealth Meteorologist. Unfortunately the data do not go back of 1908.

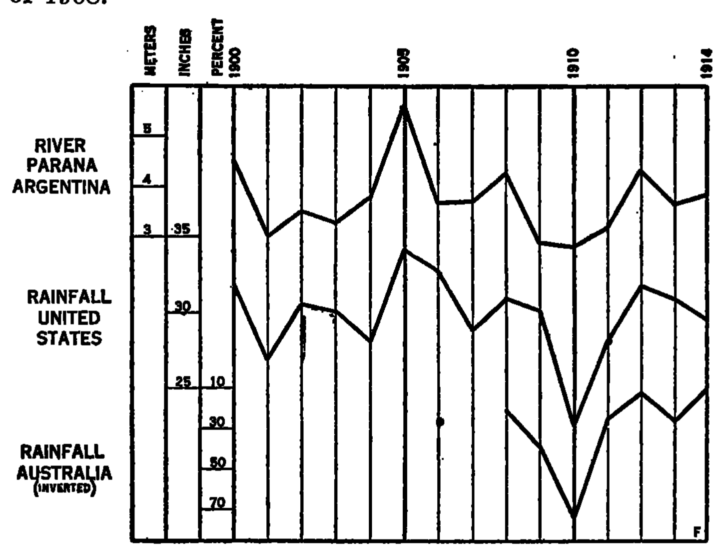


FIG. 1.—Argentine rainfall, as represented by mean annual stages of the Paraná, compared with central United States and Australian rainfalls.

Computing the correlation factor between the departures of annual rainfall in the United States and the departures of the mean river heights at Rosario by the formulas given by Yule in his Theory of Statistics, the correlation is found to be 0.71. The data are plotted in figure 1 and show to the eye the closeness of the correla-